

AD \_\_\_\_\_

Award Number: W81XWH-11-1-0582

TITLE: Progression of structural change in the breast cancer genome

PRINCIPAL INVESTIGATOR: Ryan Hartmaier

CONTRACTING ORGANIZATION: University of Pittsburgh  
Pittsburgh, PA 15213

REPORT DATE: August 2012

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

<b>REPORT DOCUMENTATION PAGE</b>				<i>Form Approved</i> <b>OMB No. 0704-0188</b>	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b></small>					
<b>1. REPORT DATE (DD-MM-YYYY)</b> 01-08-2012		<b>2. REPORT TYPE</b> Annual Summary		<b>3. DATES COVERED (From - To)</b> 1 Aug 2011 - 31 Jul 2012	
<b>4. TITLE AND SUBTITLE</b> Progression of structural change in the breast cancer genome				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> W81XWH-11-1-0582	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Ryan Hartmaier  E-Mail: hartmaier@upmc.edu				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> University of Pittsburgh Pittsburgh, PA 15213				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for Public Release; Distribution Unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> Abstract on next page.					
<b>15. SUBJECT TERMS</b> Subject terms on next page.					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			USAMRMC
			UU	19	<b>19b. TELEPHONE NUMBER (include area code)</b>

**14. ABSTRACT**

As our basic understanding of the human genome evolves, we are beginning to appreciate that it is not a static entity but rather a plastic one acquiring *de novo* mutations and structural changes. A number of recent studies suggest that breast cancer is initiated through disrupted DNA repair processes, leading to a destabilized genome, in turn promoting a heterogeneous primary lesion from which a/many subpopulation(s) acquire general or organ specific metastatic potential. I aim to identify and characterize the specific mutations that are acquired during breast cancer metastasis. To do this paired primary and metastatic breast cancer samples have been obtained and used for targeted and genome-wide analyses. Large insert mate-pair sequencing will commence in the coming months and will represent a wealth of data and will surely provide valuable results describing the process of breast cancer metastasis. Additionally, a homozygous deletion in NCOR2/SMRT was detected and is being further characterized and validated. Attached herein, I provide a detailed progress report for this project.

**15. SUBJECT TERMS**

breast cancer structural change, copy number variant

**Award:** W81XWH-11-1-0582

**PI:** Hartmaier RJ

**Report:** Annual Report (1 Aug 2011 – 31 July 2012)

**Table of Contents**

Introduction .....4

Body .....4

    Task 1: Gain necessary approvals and receive tissue samples needed for the study.....4

    Task 2: Determine impact of NCOR2/SMRT CNV on breast cancer progression .....6

    Task 3: Identification of genomic aberrations during breast cancer metastasis ..... 11

Key Research Accomplishments..... 17

Reportable Outcomes ..... 17

Conclusion..... 17

References..... 17

Appendices..... 17

Supporting Data ..... 17

## Introduction

Genomic instability is an “enabling characteristic” of cancer allowing for the acquisition of mutant, cancer-promoting phenotypes (i.e. sustained growth signaling, activation of invasion/metastasis). Large-scale cancer sequencing studies, such as The Cancer Genome Atlas (TCGA), provide an excellent resource to identify genetic events that are driving cancer. However, many passenger, non-driving events are also identified. To help interpret these results, huge sample numbers and/or complicated pathway prediction models are required. Complicating this analysis further, we are beginning to appreciate the extensive genetic heterogeneity within a tumor, like much a result from independent passenger mutations occurring throughout the evolution of the primary tumor. The approach I have proposed in this fellowship is to leverage paired tumor samples from the same patient to uncover events that have been sustained or uniquely acquired during metastatic progression. Here I describe my progress in obtaining the appropriate tissues, characterizing a candidate copy-number variation, and conducting genome-wide rearrangement detection.

## Body

*Task 1: Gain necessary approvals and receive tissue samples needed for the study (months 1-6)*

*1a: Panel paraffin embedded blocks of breast cancer progression (whole blood, DCIS, ductal carcinoma, metastatic lesion) (20 samples from each tumor for pilot study which will determine numbers needed for full study)*

The necessary approvals have been received (PRO11060660, PRO12010195). The progression samples have not been obtained as yet as the full characterization of the NCOR2 CNV has encountered some delays (see task 2). The samples for the pilot study (20 samples from each tissue type) will be obtained once the NCOR2 CNV is confidently characterized. Germline DNA was obtained from 8 individuals identified in published studies to harbor various copy number variations in NCOR2 (NA18969, NA06985, NA12044, NA12156, NA15510, NA18916, NA12248, NA18542) as well as a human variation panel of 24 individuals (HD24EC).

*1b: Matched blood, primary breast cancer, and metastatic lesions (3 tissues from 10 individuals)*

All necessary local IRB approvals for this study have been received (0506140, PRO11060660). Frozen tissues from matched normal, primary breast cancer, and metastatic breast cancer have been obtained as described in Table 1. The current state of analysis on these samples is also summarized in this table. An additional 229 paired breast cancer metastatic samples have also been identified and are in process of being obtained.

**Table 1: Summary of paired, frozen breast cancer tissues and current status of analysis.** Analyses performed include: Affymetrix Genome Wide SNP Array 6.0 (Affy6.0; genome-wide copy number), Ion Torrent Ampliseq 2.0 beta (Ampliseq2.0; targeted mutations), NanoString Copy Number Variation beta (NanoString CNV; targeted copy number), bisulfite converted RainDance targeted amplification (RainDance; targeted methylation), and large insert mate-pair sequencing (Mate-Pair; genome-wide rearrangements, copy number, and mutation). IP=in process, To send=will begin in ~1 month.

Patient	Sample Type	Site	Tumor Type	Affy6.0	AmpliSeq2.0	NanoString CNV	RainDance	Mate-Pair
RJH-MET-1	Tumor	Breast	Primary					
	Tumor	Breast to Lymph Node	Metastatic					
RJH-MET-2	Normal	Breast						To send
	Tumor	Breast	Primary					To send
RJH-MET-3	Tumor	Breast to Lymph Node	Metastatic					To send
	Normal	Buffy Coat			Yes	Yes		To send
RJH-MET-4	Normal	Spleen		Yes				
	Normal	Lt Breast					IP	
	Normal	Rt Breast						
	Tumor	Rt Breast	Primary	Yes	Yes	Yes	IP	To send
	Tumor	Lt Breast	Local Recurrence	Yes	Yes		IP	To send
	Tumor	Lt Breast	Local Recurrence				IP	
	Tumor	Breast to Liver	Metastatic	Yes	Yes	Yes	IP	To send
	Tumor	Breast to Thoracic bone	Metastatic	Yes	Yes	Yes		
	Normal	Rt Occipital		Yes	Yes			
	Normal	Rt Occipital						
	Normal	Lt Breast					IP	
	Normal	RLL Lung						
	Normal	Lt Breast						To send
	Normal	Liver						
	Tumor	Lt Breast	Local Recurrence	Yes	Yes		IP	To send
	Tumor	Lt Breast	Local Recurrence				IP	To send
	Tumor	Lt Breast	Local Recurrence					
	Tumor	Breast to Lymph Node	Metastatic	Yes	Yes		IP	To send
	Tumor	Breast to Liver	Metastatic	Yes	Yes		IP	To send
	Tumor	Breast to Rt Occipital	Metastatic	Yes	Yes		IP	To send

*Task 2: Determine impact of NCOR2/SMRT CNV on breast cancer progression (months 1-24)*

*2b: Better determine the region the CNV encompasses in lymphoblastoid cell lines and breast tumors previously identified to harbor CNV (months 1-3, samples have already been approved for use)*

Previously published data and my own preliminary data suggested that a germline CNV exists in the NCOR2 locus. This includes a number of SNP array studies, fosmid mate-pair end-sequence profiling (ESP), and QPCR based copy number analysis (Figure 1). The majority of these studies show a deletion (Figure 1 top panel, red bars), including the sequencing study, although there are a number of amplifications (Figure 1 top panel, blue bars) as well. My QPCR based copy number analysis suggested a copy number gain at the examined locus in 1/157 apparently normal individuals and also in 5/16 breast tumor samples. Of note, this assay (purchased from Applied Biosystems) utilizes an endogenous control in one region of the genome thought to be copy number neutral (RNaseP).

To better determine the extent of the copy number variation and to identify additional samples with the change, I obtained DNA from 8 individuals identified in previously published reports to harbor a copy number change in NCOR2 and a panel of 24 additional individuals and tested the same 2 previously used assays plus an additional 2 QPCR assays. Although one assay showed an amplification with the assay used in the preliminary data (individual: NA10843, assay: Hs\_03833879), the individual previously identified to harbor an amplification (NA18542) showed 2 copies in all tested regions of NCOR2. Additionally, none of the additional “positive controls” recapitulated the reported data. However, these controls were not ideal as 6 were reported to have only small deletions (<2kb) and one a large amplification (in addition to the one sample identified in the preliminary data). The samples identified to contain the large deletions (Figure 1 top panel and Figure 2) were not available from the researchers for verification. Fortunately, DNA from NA19240 (deletion identified in Kidd et al.<sup>1</sup>) is available and will be obtained as a more appropriate control for future studies.

A possible explanation of the lack of reproducibility in the QPCR CNV data is the normalization to RNaseP. Although RNaseP is likely copy number neutral in these germline DNA samples, after careful examination of the data, it appears that there may have been a technical problem with the RNaseP probe in each of the samples showing copy number gains. Additionally, it is less likely that RNaseP is copy number neutral in tumor samples. For these reasons, I utilized another strategy, the QBiomarker CNV array (QIAGEN), to validate the NCOR2 copy number change. Importantly, instead of RNaseP to normalize the data, this assay utilizes a ‘multi-copy reference’ (Mref) that recognizes a stable sequence that is repeated >40 times throughout the genome. The rationale being that if any one of the targets is altered, it will have a negligible effect on the normalization and the correct copy number can still be correctly determined. 11 assays were chosen spanning the NCOR2 locus as shown in Figure 3. These assays were run on the same 32 samples run previously. Although the same issues with the positive controls are present in this assay, interestingly one of the assays shows a homozygous deletion in 4 of the individuals and a heterozygous deletion in another 2 individuals. Further, the assay indicating a deletion the closest assay to the deletion identified by Kidd et al.<sup>1</sup> (~3kb away). Due to the difference in normalization, assay technology, and the high confidence that the copy number calls differ from 2 (all 4 samples have p-values<10<sup>-10</sup>), I feel confident that this is a real effect. Regardless, an appropriate control is essential and DNA from NA19240 will be obtained. Also, an additional assay overlapping with the Kidd et al. deletion will be obtained. Importantly, this deletion is predicted to lead to a premature stop codon, likely completely disrupting the protein expression. Associated lymphoblastoid cell lines are available for each of the tested DNA samples so this can be tested directly. Finally, since the deletion is relatively well defined (based on the Kidd et al. results), I will attempt to PCR across and sequence the breakpoint to definitively prove the structural alteration.

*2b: Develop and test FISH probes to detect SMRT CNV (months 4-6)*

Not started until copy number variation is better determined. If a PCR based test can be generated (see task 2a), this will replace the FISH probe method since it will be faster and cheaper.

*2c: Conduct SMRT CNV FISH in pilot breast cancer progression samples (months 7-9)*

Not started until copy number region is better determined.

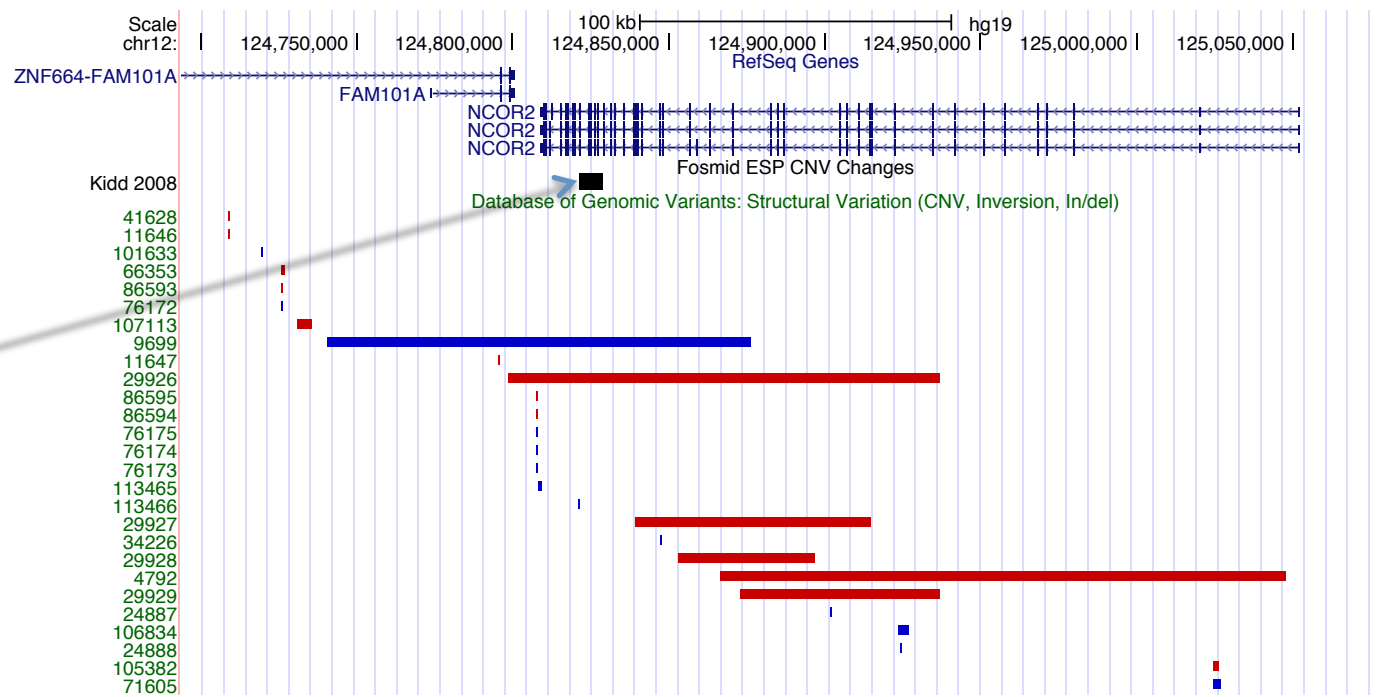
*2d: Conduct full SMRT CNV FISH study (months 10-18)*

Not started.

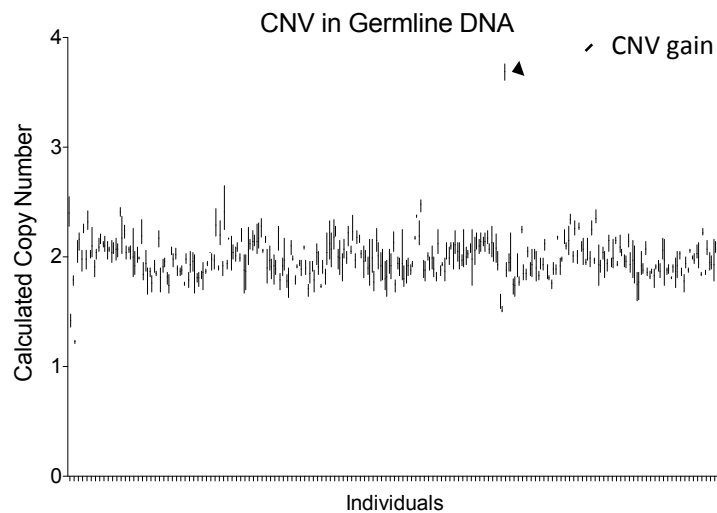
*2e: Expose CNV-negative lymphoblastoid cells (previously obtained and approved) to ionizing radiation and test for SMRT CNV acquisition (months 19-24)*

Not started

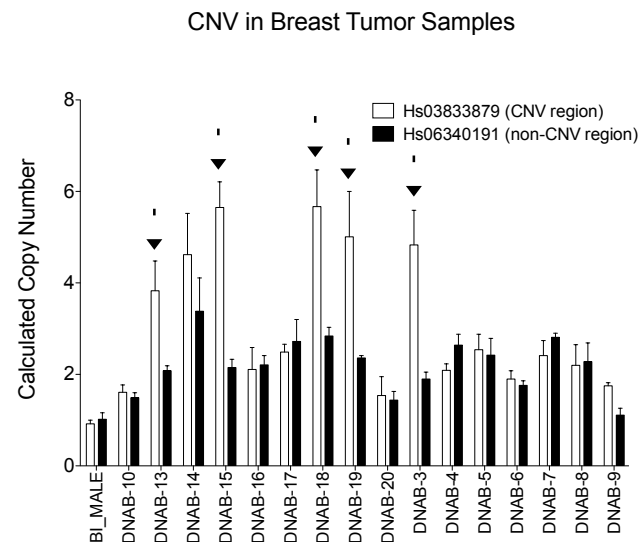
NA19240  
(deletion)

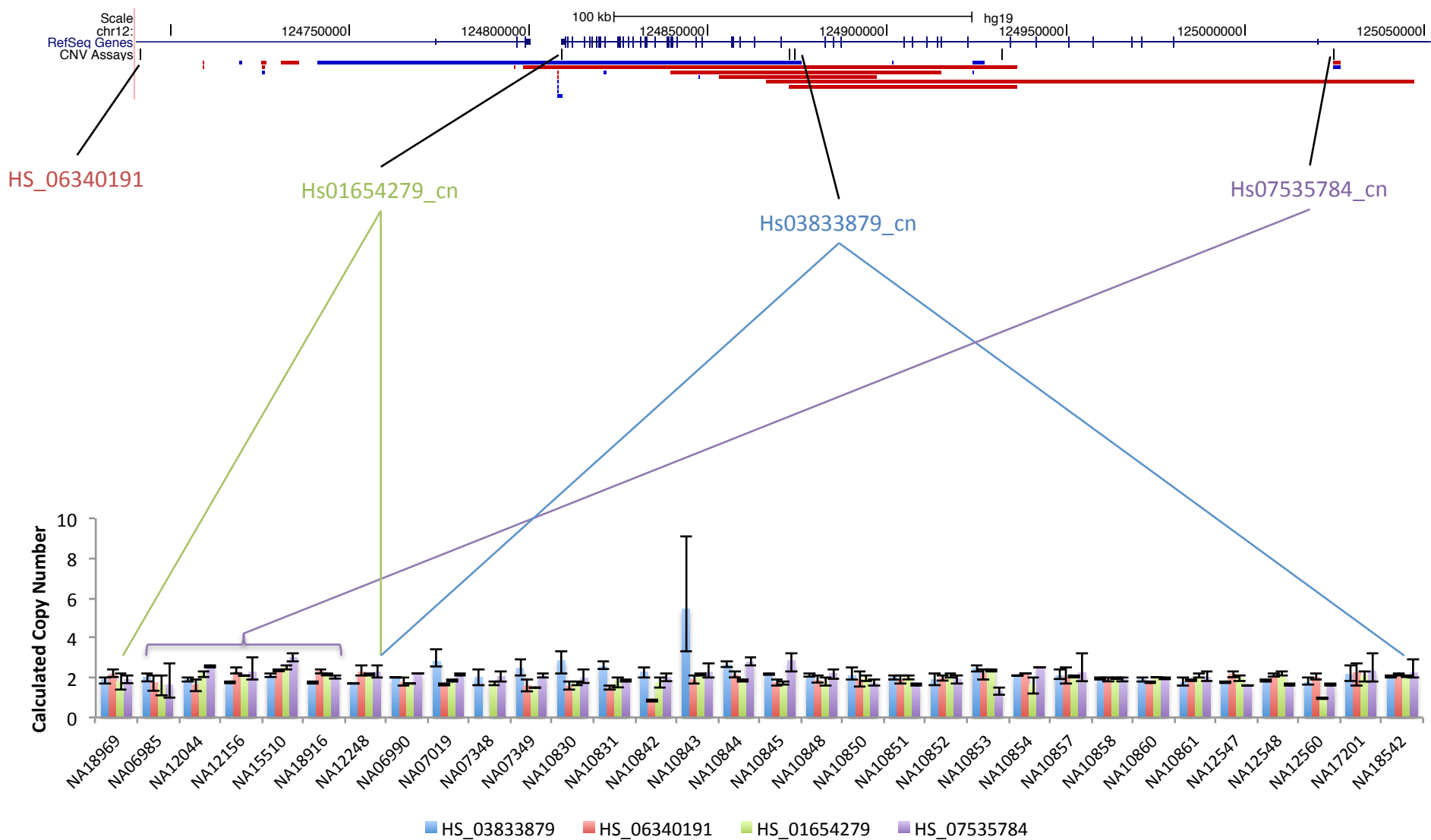


A

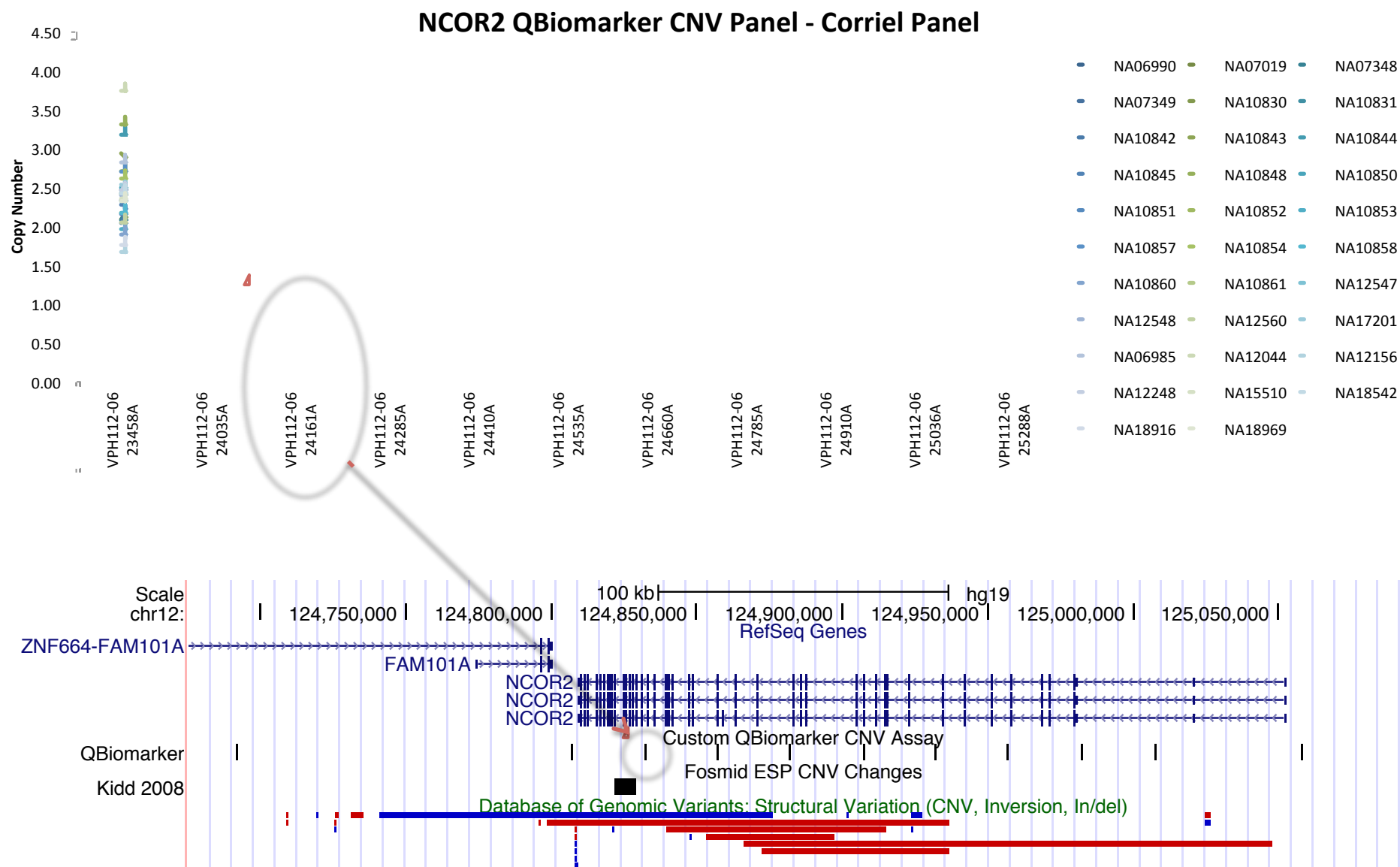


B





**Figure 2: Applied Biosystems CNV assays did not recapitulate previous NCOR2 copy number results.** The same two assays (HS03833879, HS06340191) and two additional assays (Hs01654279, Hs07535784) were run on 8 samples previously identified to harbor a copy number change in NCOR2 (highlighted in figure) and an additional 24 DNA samples from apparently normal individuals.



**Figure 3: NCOR2 QBiomarker CNV panel using multi-copy reference (Mref) identifies a deletion in NCOR2.** 11 assays + a Mref control were run on the same 32 DNA samples as in **Figure 2**. For one assay (highlighted), two DNA samples (NA18969, NA10857) were found to have a homozygous deletion while another two (NA12156, NA12547) were found to have a heterozygous deletion.

**Task 3: Identification of genomic aberrations during breast cancer metastasis (months 6-36)**

**3a: Isolate DNA from tissues and perform library preparation (3 tissues from 10 individuals) (months 6-9)**

We identified and obtained primary tumor metastasis paired tissues as described in Table 1. Two of the individuals from whom we received tumor tissue were enrolled in the Rapid Autopsy Program (RAP) at the University of Pittsburgh. A brief summary of their relevant clinical histories and available tissues is illustrated in Figure 4.

**3b: Conduct sequencing (months 10-18)**

Before conducting the proposed sequencing study, we ran a subset of the samples on a high density SNP-chip to examine if any copy number changes differed during metastasis (Affymetrix 6.0). This is a standard and well-accepted assay and will also serve as a baseline from which comparisons to future sequencing and other data types can be made.

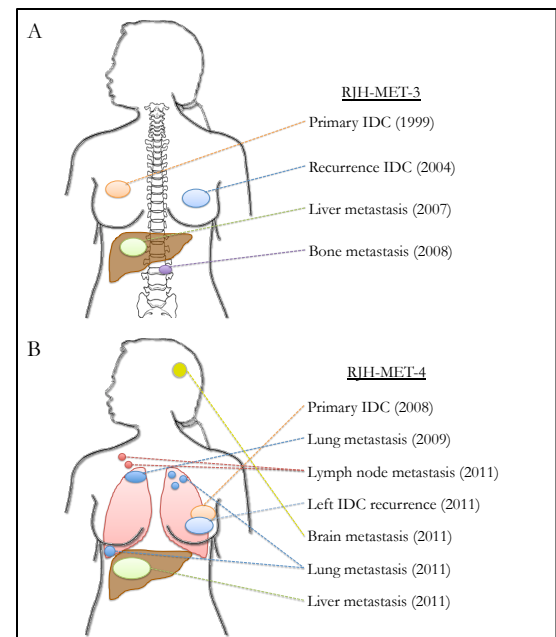
A global overview of the copy number changes is shown in Figure 5. Amplifications in chromosome 8 are seen in samples from both individuals. Interestingly, these changes seem to be muted in the distant metastatic samples, potentially reflecting a different population entering the distant metastasis. One of the most dramatic amplifications in RJH-MET-3 is in 8p11, a previously described amplification containing *FGFR1* and *WHSC1L1* (Figure 6). In RJH-MET-4, an amplification was identified in a region containing *miR-215*, a microRNA previously associated with poor prognosis and chemoresistance in colon cancer.<sup>2,3</sup> These data indicate that there are differences in the structure of the breast cancer genome during progression/metastasis.

In addition to the genome-wide copy number analysis, a targeted approach was undertaken using the NanoString Copy number panel (Figure 8A). Two cell lines (MCF-7 and HCC1954) were run as positive controls since they harbor known copy number changes. All known copy number alterations in these cell lines assayed were detected. To test the agreement of the NanoString, all copy number variable regions were intersected with regions assayed by the NanoString and their log2 copy number ratios were plotted against each other (Figure 8B). The RJH-MET-3 primary tumor and liver metastasis samples correlated well ( $r^2=0.8$  and  $0.6$  respectively) indicating that the two assays largely agreed with each other. The bone metastasis sample, however, did not correlate. Since DNA extraction from bone is extremely difficult and often results in poor quality DNA, it is likely that there is a problem with this sample and the results cannot be interpreted. Additional quality assessment will be conducted on this sample to see if it can be used in future sequencing studies.

Targeted sequencing using the Ion Torrent AmpliSeq 2.0 beta panel was conducted (Figure 8C). Again, cell line DNA was used as a positive control (MCF-7, HCC1954, MDA-MB-231, and MDA-MB-361) and all previously reported mutations assayed were identified. A known deleterious mutation in *TP53* was identified in all samples from RJH-MET-4 except the brain metastasis. Interestingly, a mutation was identified in *FGFR2* in RJH-MET-3 samples that also contain an *FGFR1* amplification (see Figure 6) indicating that FGF signaling may be important in this tumor's initiation and/or progression. Further, a novel mutation in *SMAD4* was also identified in all RJH-MET-3 samples indicating that this may be a critical mutation in this patient's disease.

Together these data indicate that novel and interesting genomic changes are present in the samples as they progress to distant metastatic disease. They also illustrate the importance and power of integration of multiple datasets (copy number with mutation data for instance). With this in mind, we are also conducting targeted methylation profiling and other studies in these matched samples to gain a complete understanding of the processes altered during metastasis. These data indicate that progressing with the sequencing study is justified. Thus, the immediate goal is to begin the large-insert mate pair sequencing in the next 1-2 months.

In preparation for the extensive data analysis that will be required for these data, I recently attended a 6-day Next Generation Sequencing workshop at the University of Pittsburgh. This was an overview covering some of the basic



**Figure 4: Clinical Histories of Rapid Autopsy patients**

tools and analyses for this data. For the more detailed training, I applied and was recently accepted to a 2-week, full-time, intensive Programming for Biologists course at Cold Spring Harbor Libraries. This course has an excellent reputation and will provide me with the necessary skills and tools in order to properly analyze my current and upcoming large datasets.

*3c: Analysis of sequencing data and basic quality control (months 19-24)*

Not started

*3d: Systematic validation of identified rearrangements (months 25-30)*

Not started

*3e: Functional studies of selected rearrangements (months 31-36)*

Not started

tools and analyses for this data. For the more detailed training, I applied and was recently accepted to a 2-week, full-time, intensive Programming for Biologists course at Cold Spring Harbor Libraries. This course has an excellent reputation and will provide me with the necessary skills and tools in order to properly analyze my current and upcoming large datasets.

*3c: Analysis of sequencing data and basic quality control (months 19-24)*

Not started

*3d: Systematic validation of identified rearrangements (months 25-30)*

Not started

*3e: Functional studies of selected rearrangements (months 31-36)*

Not started

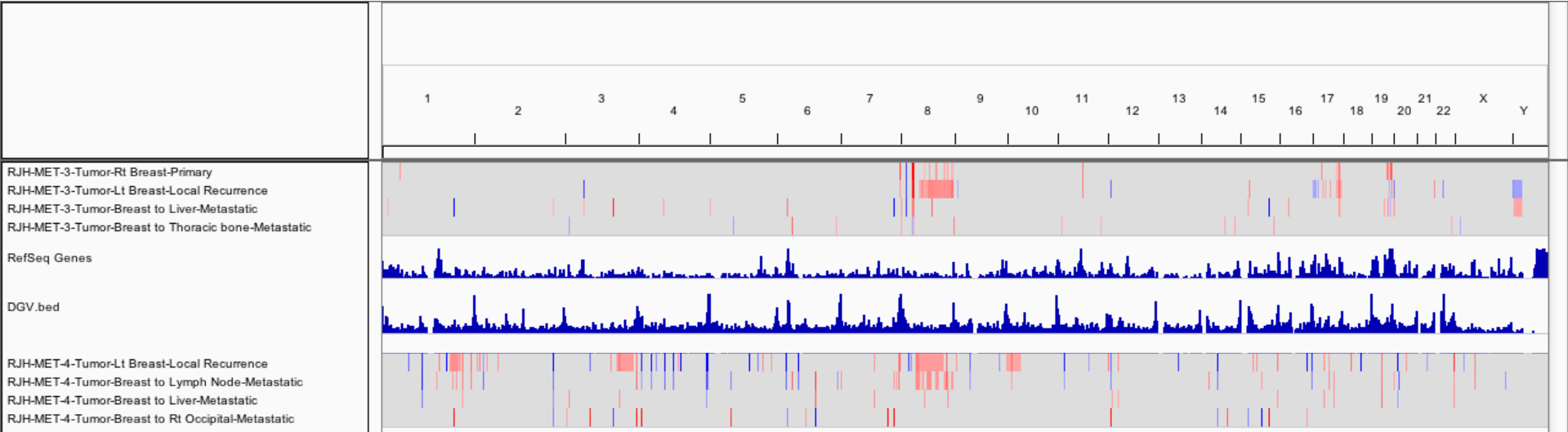
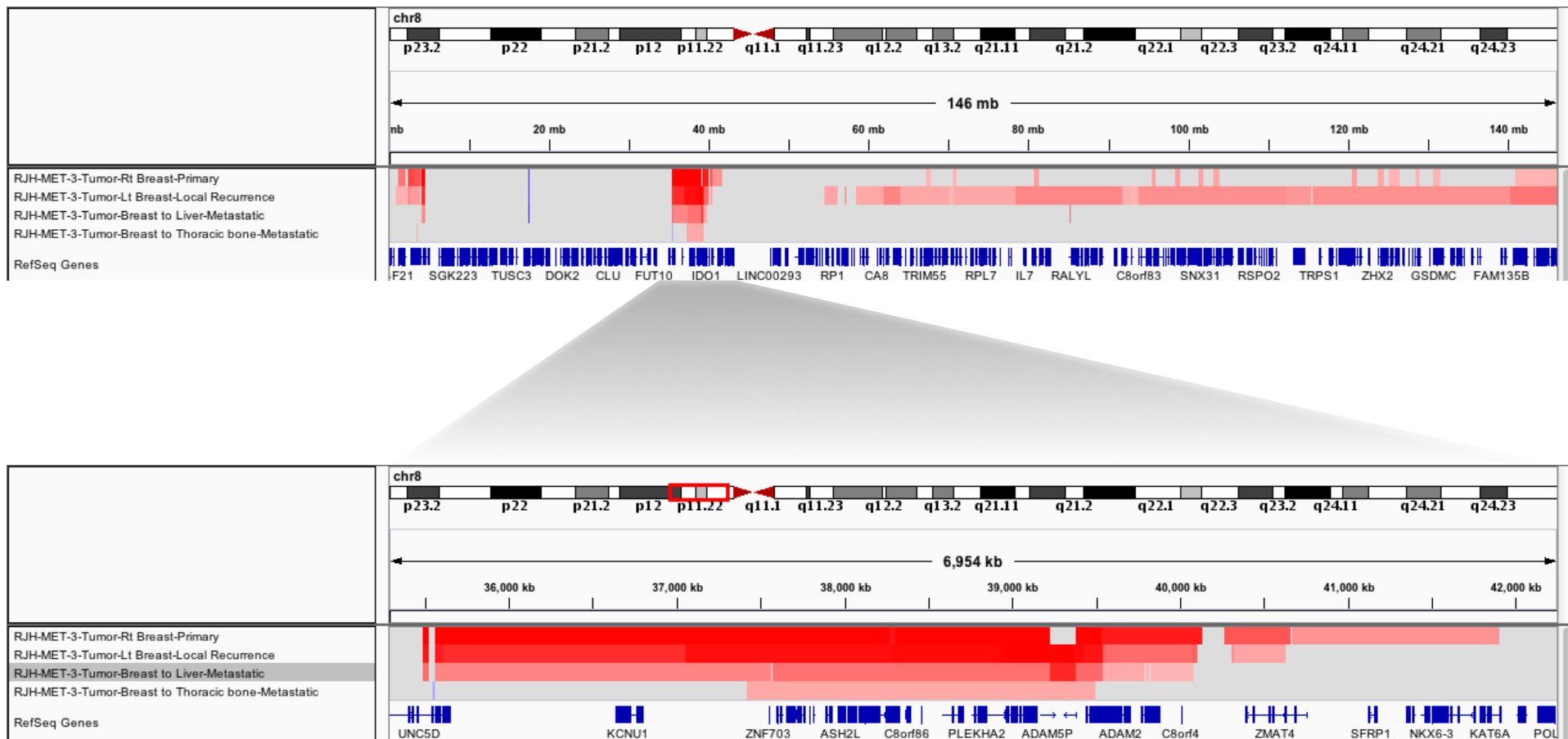
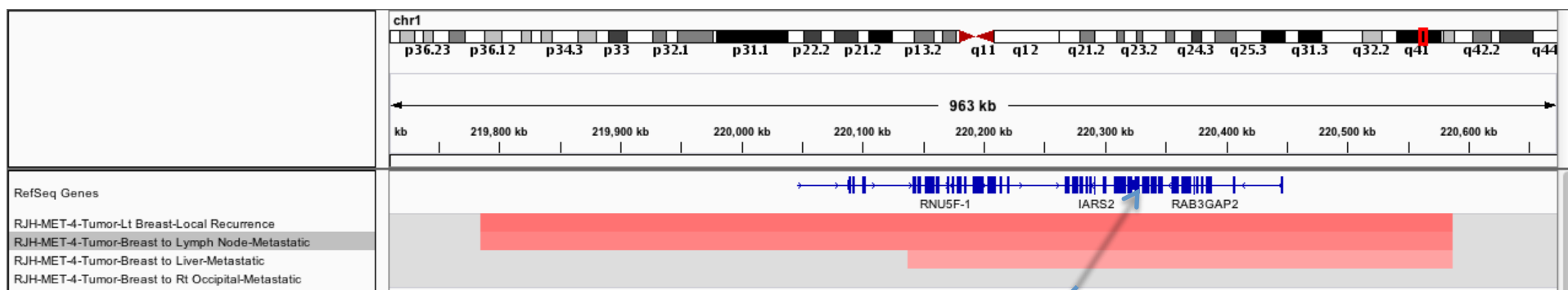


Figure 5: Overview of global copy number changes in RJH-MET-3 and RJH-MET-4 by Affymetrix 6.0. Red=amplification, blue=deletion.

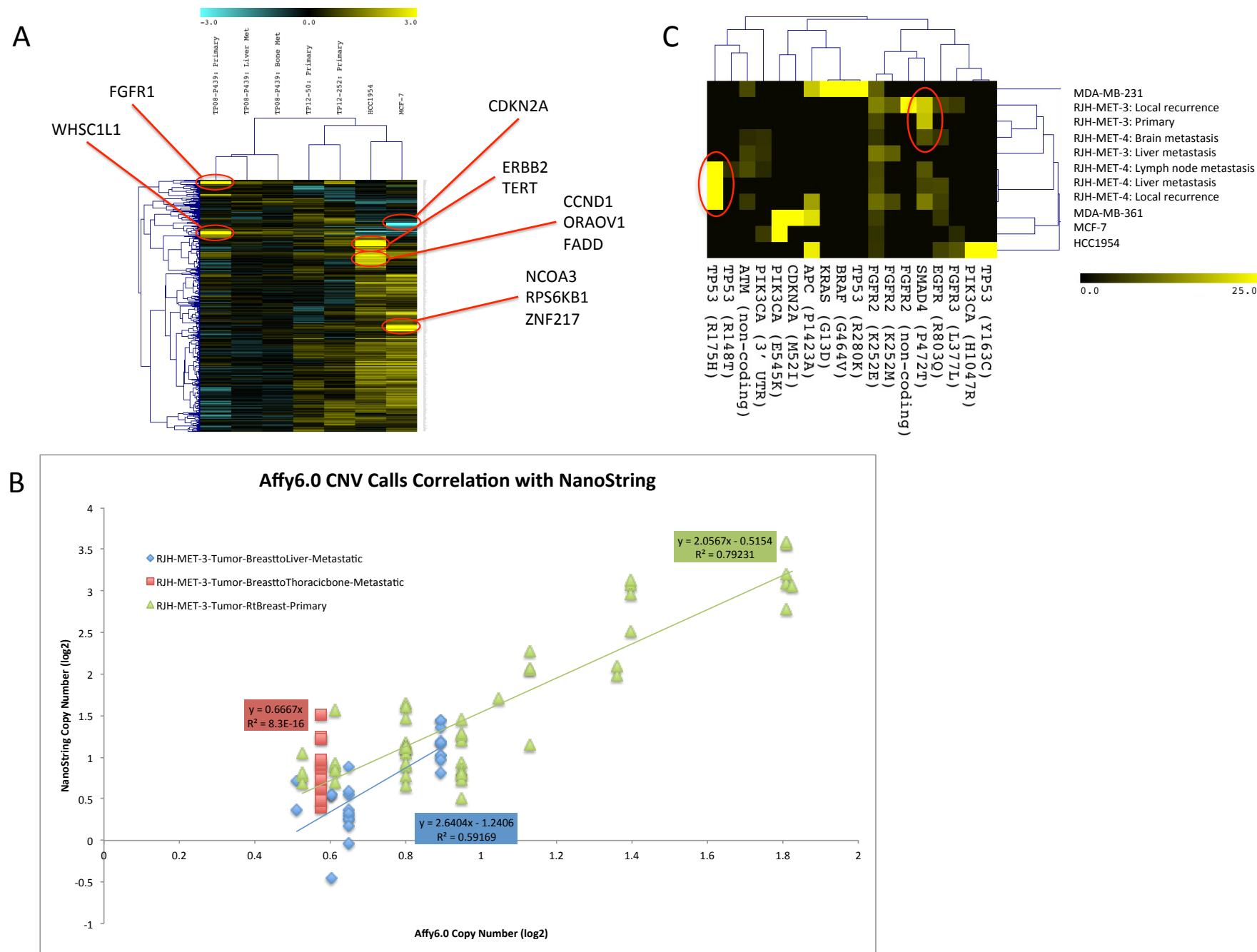


**Figure 6: Chromosome 8 copy number changes in RJH-MET-3.** Top panel shows copy number changes across chromosome 8, bottom panel shows a zoom in on 8p11 a common site of amplification in breast cancer containing *FGFR1* and *WHSC1L1*.



MIR-215

Figure 7: Copy number amplification in RJIH-MET-4 in 8q41 (miR-215).



**Figure 8: Integration of mutiple datasets on rapid autopsy samples.** (A) NanoString CNV results represented as a log2 transformed heatmap. Copy number changes highlighted for cell lines are all previously described (positive controls). (B) Copy number variable regions defined by Affy6.0 correlate well with NanoString calls for primary tumor and liver metastasis samples but not bone metastasis sample.

## Key Research Accomplishments

- Obtained matched normal-primary-metastatic tissue pairs
- Identification of a likely small homozygous germline deletion in apparently normal individuals
- Genome-wide copy number profiling on matched primary-metastatic breast tumor samples (Affy6.0)
- Validation of copy number calls via an independent method (NanoString)
- Ultra-deep sequencing of cancer related genes in paired primary-metastasis breast tumor samples (Ion Torrent)

## Reportable Outcomes

- Abstracts/Poster presentations
  - 2011.10: AACR Advances in breast cancer research
  - 2012.06: University of Pittsburgh Cancer Institute retreat
- Acceptance to Cold Spring Harbor Laboratory Programming for Biologists

## Conclusion

Since NCOR2 has a known role in tamoxifen action, the identification of the NCOR2 homozygous deletion, if completely validated, would represent a paradigm shift since a germline copy number variant could impact adjuvant cancer therapy. The genome-wide and targeted examination of acquired events during breast cancer metastasis, although just preliminary at this point, has already identified a number of interesting regions acquired or shared between primary and metastatic breast cancer metastasis. This effort is critical in our understanding of how breast cancer metastasizes.

## References

1. Kidd, J. *et al.* Mapping and sequencing of structural variation from eight human genomes. *Nature* **453**, 56–64 (2008).
2. Karaayvaz, M. *et al.* Prognostic Significance of miR-215 in Colon Cancer. *Clinical Colorectal Cancer* **10**, 340–347 (2011).
3. Song, B. *et al.* Molecular mechanism of chemoresistance by miR-215 in osteosarcoma and colon cancer cells. *Mol Cancer* **9**, 96 (2010).

## Appendices

None

## Supporting Data

None